(2)

A MODEL OF THE SUPPLY OF TRAINING IN INDUSTRY

by
L. Epstein

Research Branch
Program Development Service
DEPARTMENT OF MANPOWER AND IMMIGRATION
CANADA
1973

A Model of the Supply of Training in Industry

by

L. Epstein

Research Branch
Program Development Service
DEPARTMENT OF MANPOWER AND IMMIGRATION
CANADA
1973

ABSTRACT

Producers are assumed to decide upon the level of training to supply to employees in the context of maximizing their expected discounted sum of profits over a planning horizon. We derive a set of equations which define implicitly the producer's optimal activities and which have the following two properties: (i) linear regression techniques may be applied to them in order to estimate the technology as well as the producers' expectations; and (ii) they impose no à priori restrictions on the elasticities of substitution of the underlying technology. We rely heavily on the theory of duality and, in particular, on the concept of a variable profit function.

In Canada, the Department of Manpower and Immigration, through the Canada Manpower Industrial Training Program, is engaged in the subsidization of on-the-job training. An accurate evaluation of the program calls for an understanding of producers' behaviour in the absence of subsidization, and of the effects of varying levels of subsidization on their behaviour. The model developed in this paper is a modest step towards developing such an understanding.

The formal model developed in this study relies heavily on the concept of a variable profit function developed by Diewert [8] in an earlier paper in this series.

This study was prepared by Larry Epstein, Research Economist with the Research Branch of the Department of Manpower and Immigration.

The views expressed are solely those of the author, and do not necessarily correspond to any policy or position of the Department of Manpower and Immigration.

K. Scott,
Projection Studies Section,
Department of Manpower and Immigration.

PAPERS IN THIS SERIES

1. DIEWERT, W.E. - Canadian Labour Markets: A Neoclassical Econometric Approach. 1968.

2. DIEWERT, W.E. - Choice on Labour Markets and the Theory of the Allocation of Time. 1971.

3. DIEWERT, W.E. - The Effects of Unionization on Wages and Employment: A General Equilibrium Analysis. 1972.

4. AHAMAD, B. & SCOTT, K.F.N. - A Note on Sensitivity Analysis in Manpower Forecasting. 1972. J. Roy. Stat. Soc., Series A, Vol. 135, Part 3, 1972.

5. DIEWERT, W.E. - Walras' Theory of Capital Formation and the Existence of a Temporary Equilibrium. 1972.

6. EPSTEIN, L. - Report on a Demographic Model to Project Population by Age, Sex, Education and Province. 1972.

7. DIEWERT, W.E. - Functional Forms for Revenue and Factor Requirements Functions. 1972. To appear Int. Eco. Rev.

8. DIEWERT, W.E. - Functional Forms for Profit and Transformation Functions. 1972. To appear J. Eco. Theory.

9. GUSSMAN, Thomas Keith - Demand for Durables, Nondurables, Services and the Supply of Labour in Canada: 1946-1969. 1972.

10. EPSTEIN, L. - Some Economic Effects of Immigration: A General Equilibrium Analysis. 1972.

11. WOODLAND, A.D. - Factor Demand Functions for Canadian Industries 1946-1969. 1972.

12. DIEWERT, W.E. - Hicks' Aggregation Theorem and the Existance of a Real Value Added Function. 1973.

13. DIEWERT, W.E. - Separability and a Generalization of the Cobb-Douglas Cost, Production and Indirect Utility Functions. 1973.

14. DIEWERT, W.E. - A Note on Aggregation and Elasticities of Substitution. March 1973.

15. EPSTEIN, L. - A Model of the Supply of Training in Industry - March 1973.

16. DIEWERT, W.E. - Applications of Duality Theory, March 1973.

17. EPSTEIN, L. - Duality and Optimal Consumption Under Uncertainty, April 1973.

18. CUMMINGS, E.D. and DIEWERT, W.E. - Manpower in a General Equilibrium Context: The Model Specification. June 1973.

Copies may be obtained by writing:

Head,
Projection Studies Section,
Department of Manpower and Immigration,
OTTAWA, Ontario.
K1A 0J9

TABLE OF CONTENTS

INTRODUCTION[1/]

Recently, much work has been done on developing functional forms for the production system (and for the consumer sector) which are general in the sense that they impose few à priori restrictions on the range of possible behaviour, apart from those which are derived from theoretical considerations.[2/] Using the duality between production functions and cost or profit functions, one may very easily derive a system of factor demand equations which are consistent with profit maximizing behaviour on the part of the producer, and which do not preclude, for example, arbitrary substitution possibilities between the factors of production.[3/] The substitution elasticities, and other parameters, may thus be subject to hypothesis testing, and a fuller knowledge of producer behaviour and the underlying technology may be obtained.[4/]

Empirically testable general models of producer and consumer behaviour have been developed by Diewert in [9] and [13]. The models of the production sector, however, are confined to a one period framework, and so fail to model adequately such problems as the investment in human capital and the formation of specific capital.[5/] It is our purpose in this paper to adapt the techniques referred to above to an intertemporal framework, so as to model the producer's decision to supply training to his employees. The formation of specific capital may be modelled in a similar fashion.

In this paper, the term "training" will denote an investment in learning new skills or in perfecting old ones. Thus, it may encompass more than what is commonly referred to as "on-the-job training".

Though some figures are available on the extent of formal training going on in industry, such as in apprenticeship programs, data on the total training costs to firms, both direct and indirect, in all types of training, are very rare. That the costs to employees are significant was shown by Mincer in [24]. He states (P.73) that even in the 1950's "investment in on-the-job training is a very large component of total investment in education in the United States economy. Measured in terms of costs, it is as important as formal education for the male labour force and amounts to more than a half of total (male and female) expeditures on school education." All indications are that the decision to supply training is an important one for the firm as well.6/

In [2] Becker discusses two types of training, general and specific. Completely general training is defined as the type of training that increases the marginal productivity of trainees by exactly the same amount in the firms providing the training as in other firms. Hence, in a competitive labour market, the wage rates of the trainees will rise by the same amount as their marginal productivity, and the firm providing the training does not capture any of the return. Therefore it will not bear any of the costs involved and will supply general training only if the trainees are willing to pay all of them. From the point of view of the firm's behaviour, this is very much a one period problem and hence may be so modelled.

We shall concern ourselves principally with specific training, in which the productivity of trainees is increased more in firms providing the training than in other firms. Consequently, firms will

benifit in the future from the excess of the marginal productivity of trainees over their wage rates,(the extent of the return to the firms depending upon the number of trainees who remain with them after completing training),and so will be induced to supply specific training. On the other hand, costs are also incurred and may include the following: direct training costs (such as payment of intructor's fees, or tuition fees at a formal institution of learning); foregone production as compared with the use of a fully qualified worker, (during the learning process the trainee probably receives a wage exceeding his marginal productivity); increased materials wastage and possible increased wear and tear on machines; and costs of supervision - either by taking a trained worker off his job, leading to foregone production, or any losses of output suffered by a general relaxation of supervision elsewhere due to attention given to the learner. The producer must decide on a trade-off between present costs and future returns.

Becker has given a set of marginal conditions to describe this trade-off decision. Upon assuming a long run equilibrium condition of zero discounted sum of profits, he shows that the usual one period equilibrium condition that wages= marginal productivity, is replaced by the condition that the sums of the discounted streams of expected wage rates and expected marginal products be equal. We extend this condition to the case of a firm that wishes to maximize the discounted (non zero) sum of the expected stream of profits over the planning horizon. We develop an empirically testable model of the producer's behaviour, including his decision to supply training.

We note that some studies have treated time as a continuous variable, so that discounted sum of expected future profits is replaced by an integral over time, and the techniques of optimal control theory are used to derive the profit maximizing activities.[7/]

However, this approach forces one to make some very restrictive assumptions about the underlying production function. As a matter of fact, we have not found in the literature any continuous time analysis of the supply of training which goes beyond a Leontief, fixed factor proportion production function.[8/] Using duality we are able to consider general production functions.

Our model may be briefly outlined as follows: A producer, facing a two period time horizon, may produce an output good via a four factor production function. The factors are capital, "unskilled" labour, "trainee" labour and "skilled" labour. The producer is aware of first period prices, all of which are determined exogenously, except those for the latter two factors where he has some monopsony power.[9/] He also forms expectations of second period (future) prices, and of the monopsony power he will possess in the second period, with the supply curve of "skilled" labour which he expects to be facing in the future being influenced by the amounts of "skilled" and "trainee" labour which he employs in the first period. In striving to maximize the sum of expected profits, discounted by the expected discount rate, the producer decides on his first period activities and on his plans for the future. These optimal activites and plans are defined by a set of equations involving the variable profit function, a dual function to the producer's production function. We hypothesize a flexible functional form, derived by Diewert in [11], for the profit function, which, when substituted into the set of equations, lends itself to simple regression techniques for the estimation of its parameters.

We note that many of the simplifying assumptions we have made, such as the number of time periods, output goods, and factors of production, may be easily removed. (Indeed, the extension of the model to several time periods is described briefly). Many more types of labour, some perhaps defined partially by the age of workers, and many types of training, may be introduced, but at the cost of increasing the complexity of the model. The investment in specific capital formation may also be modelled.[10/]

In the next section, we review some preliminary material from duality in the theory of production, with which the reader may not be too familiar. We then proceed to present our model in more detail. In the final section we comment briefly on how our model may be applied empirically.

I. SOME RESULTS IN THE THEORY OF DUALITY [11/]

The usual starting points of the theory of production have been: (i) the production possibilities set, which is the set of all input and output combinations which the firm can produce with available technology;or equivalently, (ii) the production function, which, for any feasible production plan, gives the maximum net output of any commodity as a function of the net outputs of all other commodities. In [26] Shephard showed that a single output multiple input technology could equally as well be described by a minimum cost function, which would give the minimum cost, necessary to produce a given net output as a function of input prices. McFadden [21], [22] and Gorman [16] have shown that a multiple output, multiple input technology may be equivalently described by a profit function, which gives the maximum profit for any feasible production plan as a function of the prices of all goods. In [11] Diewert has established a duality between production possibilities sets, production or transformation function, and variable profit functions, which, apart from fixed costs, give the maximum profit attainable, holding some inputs fixed, as a function of the prices of the remaining commodities. For a more detailed discussion and proofs of these results, the reader is asked to refer to Diewert's paper and to the other references. We shall now state more precisely the results proved in [11] and which we shall need for our model.[12/]

We consider a technology with I+J inputs and outputs. Denote the variable inputs and outputs by the I dimensional vector

$u=(u_1,\ldots,u_I)$, and the fixed inputs by the J dimensional vector $v=(v_1,\ldots,v_J)$. Outputs are indexed positively and inputs negatively. The production possibilities set T is defined by the following conditions:

(i) T is a closed, nonempty subset of I+J dimensional space;

(ii) if $(u;v)\in T$, then $v\leq o$ (last J goods are always inputs);

(iii) T is a convex set (non-increasing marginal rates of transformation);[13/]

(iv) T is a cone (constant returns to scale);[14/]

(v) if $z\in T$ and $z'\leq z$, then $z'\in T$ (free disposal);

(vi) if $(u;v)\in T$, then the components of u are bounded from above (for finite fixed inputs, the set of producible outputs is bounded).

An alternate approach to production theory is via the notion of a production or transformation function t, which is defined by the following conditions:

(i) t is an extended real valued function defined and bounded from above for each I-1+J dimensional vector $(\hat{u};v)$, such that $v\leq o$, (where $\hat{u}$ is an (I-1) dimensional vector), satisfying $t(0;0)=0$.

(ii) t is a continuous from above function;[15/]

(iii) t is a (proper) concave function;[16/]

(iv) t is positive linear homogeneous in $\hat{u}$ and v, i.e., $t(\lambda\hat{u};\lambda v)=\lambda t(\hat{u};v)$ for every scalar $\lambda>0$;

(v) t is nonincreasing in the components of $(\hat{u};v)$;

(vi) for every $v\leq o$, the set $\{(u_1,\hat{u}) \mid u_1\leq t(\hat{u};v)\}$ is bounded from above.

Given a production possibilities set T satisfying conditions A, one may define a transformation function satisfying conditions B, as follows:

I.1 $$t(\hat{u};v)=\begin{cases}\max\limits_{u_1}\{u_1 \mid (u_1,\hat{u};v)\in T\} & \text{if there exists a } u_1 \text{ such that } (u_1,\hat{u};v)\in T\\ -\infty & \text{otherwise,}\end{cases}$$

for all vectors $\hat{u}=(u_2,u_3,---u_1)$ and $v\leq 0$.

On the other hand, given a tranformation function t, the corresponding production possibilities set may be obtained as follows:

I.2 $$T=\{(u_1,\hat{u};v) \mid u_1\leq t(\hat{u};v),\ v\leq 0\}.$$

So defined, T satisfies conditions A. In addition the transformation function corresponding to T as defined by I.1 is the function t with which we started. That is, these are equivalent approaches to production theory.

Given a production possibilities set T and supposing that the producer's fixed inputs are fixed at v, and that he can buy or sell variable inputs or outputs at the fixed positive prices $p=(p_1,p_2,\ldots,p_J)>>0$ the variable profit function may be defined as follows:

I.3 $$\pi(p;v)\equiv\max_u\{p^Tu \mid (u;v)\in T\},\text{ where } p>>0 \text{ and } v\leq 0.$$

We can show that the variable profit function so defined satisfies the following set of conditions:

C. (i) $\pi(p;v)$ is nonnegative and bounded above by p^Tb for a fixed vector b if $p>>0$, $v\leq 0$ and v is bounded from below;

(ii) π is positive linear homogeneous in p, i.e., for every $\lambda\geq 0$, $\pi(\lambda p;v)=\lambda\pi(p;v)$;

(iii) π is convex and continuous in p for every fixed $v \leq o$;[17]

(iv) π is non-increasing in v for every fixed p;

(v) π is concave and continuous in v for every fixed p;

(vi) π is positive linear homogeneous in v, i.e. for every $\lambda \geq o$, $\pi(p;\lambda v) = \lambda\pi(p;v)$.

Conversely, given a function π satisfying these conditions, the corresponding production possibilities set is defined by:

I.4 $$T \equiv \{ (u;v) \mid p^T u \leq \pi(p;v) \text{ for every } p \gg 0 \text{ and } v \leq 0 \}.$$

So defined, T satisfies conditions A. Moreover, the profit function corresponding to it via definition I.3 is precisely the function π with which we began; that is, we have three equivalent approaches to production theory.

The following lemma is of prime importance.

Modified Hotelling Lemma:[18]

If a variable profit function $\pi(p;v)$ satisfies conditions C and is in addition differentiable with respect to the prices of variable outputs and inputs at $p^* \gg 0$, $v^* \leq 0$, then we have $\frac{\partial\pi}{\partial p_i}(p^*;v^*) = u_i(p^*;v^*)$, $i = 1,2,\text{---}, I$, where u_i is the profit maximizing amount of output i (of input i if $u_i(p^*;v^*) < 0$) given prices p^* and fixed inputs v^*.

From an econometric point of view, this lemma is very useful, for it enables us to obtain functional forms for demand and supply functions consistent with profit maximization simply by choosing a suitable functional form for π and differentiating it with respect to input and output prices.

Diewert has hypothesized the following functional form for π:

I.5 $$\pi(p;v)=\sum_{i=1}^{I}\sum_{h=1}^{I}\sum_{j=1}^{J}a_{ih}\left(\tfrac{1}{2}p_i^2+\tfrac{1}{2}p_h^2\right)^{\frac{1}{2}}x_j+\sum_{i=1}^{I}\sum_{j=1}^{J}c_{ij}p_ix_j+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{J}b_{jk}p_ix_j^{\frac{1}{2}}x_k^{\frac{1}{2}},$$

where $a_{ih}=a_{hi}$ $\quad 1\le h<i\le I$

$b_{jk}=b_{kj}$ $\quad 1\le k<j\le J$

$a_{ii}=0,$ $\quad 1\le i\le I$, $\quad b_{jj}=0$ for $1\le j\le J$,

and where $x=-v$.

For many choices of the set of parameters, this function satisfies the regularity conditions C over some nonempty set of vectors $(p;v)$, and so is "locally" a variable profit function.[19/] In addition the functional form is flexible in that the parameters may be chosen so as to provide a good local approximation to an arbitrarily given twice differentiable profit function.

Lemma:

Let π^* be an arbitrary variable profit function (consistent with conditions C) which is twice continuously differentiable at the point $(p;v)$ where $p>>0, v<<0$. Then, in general, there exists a $\pi(p;v)$ given by I.5 which provides a second order approximation to π^* at the point $(p;v)$.

The importance of this lemma lies in the fact that the first and second order partial derivatives of a variable profit function completely determine the elasticities of substitution of the underlying technology.[20/]

The functional form in I.5 possesses two additional advantages. Firstly, note that the parameters appear linearly in the defining equation. Hence, they will appear linearly also in the equations defining the partial derivatives $\partial\pi(p;v)/\partial p_i$ $i=1,2,---,I$, (which by

the Modified Hotelling Lemma are the variable output supply and input demand functions), and $\partial\pi(p;v)/_{\partial v_j}$, j=1,2,---J. This will enable us to apply linear regression techniques to estimate the parameters of our model.

Secondly, à priori expectations that, for example, the variable profit maximizing amount of variable output (or input) i is not responsive to a change in the price of the variable output (or input) h, may be translated into the restriction that $a_{ih}=0$.[21/] By setting $b_{jk}=0$, we may impose à priori the restriction that the marginal (variable) profit of the j^{th} fixed input (or output) is not responsive to a change in the level of the k^{th} fixed input (or output).[22/] Such à priori restrictions on the parameters may be helpful in the event that, because of a small number of observations, we are not able to estimate all of the parameters.

II. THE MODEL

Consider a producer who faces a time horizon of two periods. On the basis of his ex ante price expectations, he chooses from among available technologies that technology which maximizes the expected sum of his discounted profits, and which is then taken to be fixed for the time horizon.[23] We represent the technology by means of a four factor constant returns to scale production function F (K, L_1, L_1^T, L_2) satisfying the regularity conditions B[24], where the input factors are capital K, "unskilled" labour L_1, "trainee" labour L_1^T, and "skilled" labour L_2.[25] "Unskilled" workers, of course, are incapable of supplying L_2, while they may enter training. Trainees gain the potential to supply L_1 in the following period. Since the training is specific, the marginal productivity of "skilled" labour in the firm providing the training is higher than the marginal productivity in other firms. As noted above, there is a future return to the firm from supplying training, and so its first period activities derive from intertemporal considerations.

In order to be more precise, denote by Y_1 and Y_2 the output produced in (or planned for) periods 1 and 2 respectively, and add a subscript to each of the input factors to indicate the time period under consideration. (For example, K_2 and L_{12} will represent the amounts of capital and of "unskilled" labour respectively used in period 2).

The first period prices p_1 (for Y_1), r, (for K_1), and w_{11} (for L_{11}), are assumed to be exogenous to the firm. We argue, however, that to some extent the producer is free to vary the equilibrium wage rate for "skilled" labour, and by so doing to influence the amount of "skilled" labour that is offered him. For certainly, since we are dealing with specific training, the equilibrium wage rate is not uniquely defined by the marginal productivity of "skilled" labour in the economy. On the other hand, the following disequilibrium situation is conceivable: if the wage rate offered "skilled" labour were sufficiently high, then all "unskilled" workers would flood into training, in anticipation of high future earnings which would more than compensate for the costs incurred in the first period in undergoing training. However, assuming that workers attempt to maximize an intertemporal utility function, distinct occupational preferences and expectations for the future ensure that the equilibrium wage may be varied at least over some interval. We assume also that an increase in the equilibrium wage in this interval will elicit an increased supply of "skilled" labour to the firm. The supply curve of "skilled" labour facing the firm in the first period, after incorporating the effects of first period marker prices, may thus be described by the following (inverted) supply of "skilled" labour function:

$$w_{21} = w_{21}\ (L_{21}; p_1, r_1, w_{11})$$ 26/

We suppose that w_{21} is a continuous function and in addition is twice continuously differentiable with respect to L_{21}. 27/

We have already specified that the supply curve be upward sloping,

i.e., that $\partial w_{21}/\partial L_{21} > 0$.

Note that because of the above argument, the function in II.1 may be defined only for L_{21} in some interval. Since all possible equilibrium wage rates correspond to values of L_{21} in that interval, however, the equilibrium supply of "skilled" labour is completely described by II.1.

The above reasoning may be applied also to justify the assumption of an (inverted) supply function of "trainee" labour given by:

I.2 $$w_{11}^{T} = w_{11}^{T}\,(L_{11}^{T}, L_{21}; p_1, r_1, w_{11}).$$ [28/]

We suppose that w_{11}^{T} is a continuous function and in addition is a twice continuously differentiable function of its "labour" arguments. Once again we assume an upward sloping supply curve, i.e., $\partial w_{11}^{T}/\partial L_{11}^{T} > 0$.

The supply of "trainee" labour, reflecting wage expectations and intertemporal utility maximization on the part of workers, will also depend on w_{21} and hence on L_{21}. One would expect that the higher the current "skilled" wage, the greater the expectations of workers for "skilled" wages in the future, and hence the more willing they will be to invest in training.[29/] Therefore, we would expect that $\partial w_{11}^{T}/\partial L_{21} \leq 0$.

A few words might be in order to justify the assumption of price taking behaviour on the part of the producer with respect to some goods and factors, and monopsony power with respect to "trainee" labour and "skilled" labour. This may be rationalized

as follows: "Unskilled" labour and capital are used in many industries and so the producer has no control over their prices. Now, we may think of our producer as using a technology distinct from that used by other producers in the same industry, and hence gaining monopsony power over "trainee" and "skilled" labour while remaining a price taker with respect to the output good. Alternately, his monopsony power may derive from spatial considerations, i.e., he may be the sole producer in that industry in a geographic region, and so, goods being much more mobile than labour, the above situation may result.[30]

The producer is aware of the supply curves II.1 and II.2 and of the first period market determined prices, and based upon this knowledge forms the following price expectations for period 2:

II.3 (expected price of Y_2) $p_2=p_2(p_1, r_1, w_{11})$

(expected price of capital) $r_2=r_2(p_1, r_1, w_{11})$

(expected wage rate of "unskilled labour) $w_{12}=w_{12}(p_1, r_1, w_{11})$

(expected discount rate between periods 1 and 2) $\rho=\rho(p_1, r_1, w_{11})$.[31]

We assume that these functions are continuous and assume finite positive values.

The producer also forms expectations of the supply curve for "skilled" labour that he will face in the future. In addition to prevailing and expected market prices, the expected supply curve will be influenced by the number of workers trained in period 1 and their expected rate of attrition from the firm. We also allow for the

effect which the number of "skilled" workers employed by the firm in period 1 might have on the anticipated supply curve. One might argue, for example, that a "skilled" worker employed in the firm in period 1 would demand a lower wage to continue working there, than would a new "skilled" worker to join the firm and abandon his first period activities. All of these influences are summarized in the expected (inverted) supply function described by:

II.4 $$w_{22}=w_{22}\ (L_{22},L_{11}^{T},L_{21};\ p_1,\ r_1,w_{11}).$$ [32/]

We assume that w_{22} is continuous in all its arguments jointly, and that it is twice continuously differentiable with respect to its first three arguments. We would expect that $\partial w_{22}/\partial L_{11}^{T}\leq 0$, i.e., the greater the amount of training supplied in period 1, the lower the wage rate which will have to be paid to secure a given supply of "skilled" labour in period 2.

Similarly, on the basis of the reasoning given above, one might expect that $\partial w_{22}/\partial L_{21}\leq 0$. However, it is far from clear that this is so in general. For if the amount of "skilled" labour employed in period 1 is increased, and hence a higher wage rate w_{21} paid, "skilled" workers may form high expectations and be more reluctant than previously to work at a given wage rate in the second period. The net effect may be to decrease the supply of "skilled" labour at a given wage. Hence the sign of $\partial w_{22}/\partial L_{21}$ is in general ambiguous.

Lastly, we suppose that the expected supply curve is upward sloping, i.e., that $\partial w_{22}/\partial L_{22}>0$.

As an example of such an expected supply function consider that function w_{22} defined implicity by the following equation:

II.5 $$L_{22}= \alpha(w_{22},L_{11}^{T},p_1,r_1,w_{11})\ L_{11}^{T} +\beta(w_{22},L_{21},p_1,r_1,w_{11})\ L_{21},$$

where α and β are increasing functions of w_{22} attaining values between 0 and 1.[33/] We will have $\partial w_{22}/\partial L_{11}^{T} \leq 0$ if $\alpha + L_{11}^{T} \cdot \partial\alpha/\partial L_{11}^{T} > 0$.[34/] Simarly, the sign of $\partial w_{22}/\partial L_{21}$ is the opposite of the sign of $\beta + L_{21}\cdot\partial\beta/\partial L_{21}$. If α and β satisfy the appropriate continuity and differentiability conditions, the w_{22} so defined satisfies all the conditions stated above. Of course, $1-\alpha$ and $1-\beta$ are the expected rates of attrition from the firm of "trainee" and "skilled" labour respectively.

For the sake of simplicity we assume that the producer supplies no training in the second period. This is not unreasonable if we consider that in a two period model the demand for training would probably be so low so as to make it unprofitable for the producer to supply and training. Hence, we do not specify an expected supply curve, w_{12}^{T}, of "trainee" labour.

Since training may involve direct costs to the producer (other than wages paid to trainees) such as fees paid to special instructors, we denote by $C(L_{11}^{T})$ the direct costs involved in supplying L_{11}^{T} units of training in period 1. We assume that C is differentiable, and that $C^1(L_{11}^{T})\geq 0$.

The basic assumption of the model is that the producer maximizes the expected discounted sum of his profits over the two periods;[35/] i.e.that he solves the following maximization problem:

II.6
$$\max_{\substack{Y_1,Y_2\\K_1,K_2\\L_{11},L_{12}\\L_{11}^T,\\L_{21},L_{22}\text{, all nonegative}}}\left\{\begin{array}{l}(p_1Y_1-r_1K_1-w_{11}L_{11}-w_{11}^TL_{11}^T-w_{21}L_{21}-C(L_{11}^T))\\ +\frac{1}{1+\rho}(p_2Y_2-r_2K_2-w_{12}L_{12}-w_{22}L_{22})\end{array}\;\middle|\;\begin{array}{l}Y_1\leq F(K_1,L_{11},L_{11}^T,L_{21})\\Y_2\leq F(K_2,L_{12},0,\;L_{22})\end{array}\right\}$$

Breaking the maximization into two stages, where the producer first maximizes with respect to those commodities whose (expected) prices are independent of his actions, and secondly with respect to the remaining commodities, yields the following equivalent problem:

II.7
$$\max_{L_{11}^T,L_{21},L_{22}\geq 0}\left\{\begin{array}{l}\max\limits_{Y_1,K_1,L_{11}\geq 0}\left\{p_1Y_1-r_1K_1-w_{11}L_{11}\;\middle|\;Y_1\leq F(K_1,L_{11},L_{11}^T,L_{21})\right\}\\ -w_{11}^TL_{11}^T-w_{21}L_{21}-C(L_{11}^T)\\ +\max\left\{\frac{p_2}{1+\rho}Y_2-\frac{r_2}{1+\rho}K_2-\frac{w_{12}}{1+\rho}L_{12}-\frac{w_{22}}{1+\rho}L_{22}\;\middle|\;Y_2\leq F(K_2,L_{12},0,L_{22})\right\}\end{array}\right\}$$

Recalling the definition of the variable profit function (I.3) and denoting by π the variable profit function corresponding to F, with fixed inputs of "trainee" and "skilled" labour, this problem may be rewritten as follows:

II.8
$$\max_{L_{11}^T,L_{21},L_{22},\geq 0}\left\{\begin{array}{l}\pi(p_1,r_1,w_{11};L_{11}^T,L_{21})-w_{11}^TL_{11}^T-w_{21}L_{21}-C(L_{11}^T)\\ +\frac{1}{1+\rho}\pi\,(p_2,r_2,w_{12};0,L_{22})-\frac{w_{22}}{1+\rho}L_{22}\end{array}\right\}$$

Implicit in the discussion to this point has been the assumption that the maxima in the above problems exist. Since there are constant returns to scale in production, we will have to stipulate conditions for the labour supply equations II.1 and II.2, and for the expectation function II.4 in order to ensure that the producer does not expect to earn indefinitely large profits over the two periods.

Consider II.8. Since the sum of the variable profit functions is linear homogeneous concave in $(L_{11}^{T}, L_{21}, L_{22})$, it is easy to see that the maximum will exist if input costs are sufficiently increasing. In particular, we require that

$$C(L_{11}^{T}) + w_{11}^{T}(L_{11}^{T}, L_{21})\, L_{11}^{T} + w_{21}(L_{21})L_{21} + w_{22}(L_{22}, L_{11}^{T}, L_{21})\, L_{22}$$

is an increasing convex function of $(L_{11}^{T}, L_{21}, L_{22})$, given the prices p_1, r_1, w_{11}.[36/]

Monotonicity is reasonable in that one would expect that the greater the level of inputs used, the greater the total cost of these inputs to the producer. (We do not bother to write out the inequalities in the first order partial derivatives of the above function necessary for monotonicity. Most of the first order partials involved have already been assumed to be nonegative). Convexity implies increasing marginal input costs for each factor, and increasing marginal rates of substitution between factors at a constant total cost level. In the event that there is strict convexity above, the maximum profit will be attained at a unique point. Making this assumption, we may proceed to find the profit maximizing activities chosen by the producer.

We make the following definition:

II.9 $$P\left(p_1,r_1,w_{11},\frac{p_2}{1+\rho},\frac{r_2}{1+\rho},\frac{w_{12}}{1+\rho};L_{11}^T,L_{21},0,L_{22}\right)\equiv \pi(p_1,r_1,w_{11};L_{11}^T,L_{21})+ \pi\left(\frac{p_2}{1+\rho},\frac{r_2}{1+\rho},\frac{w_{12}}{1+\rho};0,L_{22}\right)$$

It is easy to see that P satisfies with respect to its arguments all the regularity conditions C, and so is a variable profit function. Indeed it is the variable profit function corresponding to our two period production possibilities set

$$\left\{(Y_1,K_1,L_{11},L_{11}^T,L_{21},Y_2,K_2,L_{12},0,L_{22}) \;\middle|\; \begin{array}{l} Y_1\le F(K_1,L_1,L_{11}^T,L_{21}) \text{, all variables } \ge 0 \\ Y_2\le F(K_2,L_2,0,\ L_{22}) \end{array}\right\}$$

Therefore, by the Modified Hotelling Lemma, the profit maximizing levels of variable inputs and outputs are obtained by differentiating P with respect to the appropriate prices.

In addition, maximizing problem II.8 becomes:

II.10 $$\max_{L_{11}^T,L_{21},L_{22}\geq 0}\left\{P - C(L_{11}^T) - w_{11}^T\ L_{11}^T - w_{21}\ L_{21} - w_{22}\ L_{22}\right\},$$

so that if the maximum is attained at an interior point, that point must satisfy the first order conditions of a stationary point. Combining these two arguments, we derive the following set of equations which define the profit maximizing set of inputs and outputs:

II.11 (i) $$\frac{\partial\pi}{\partial p_1}(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^*) = Y_1^*$$

(ii) $$\frac{\partial\pi}{\partial r_1}(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^*) = K_1^*$$

(iii) $$\frac{\partial\pi}{\partial w_{11}}(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^*) = L_{11}^*$$

(iv) $\frac{\partial\pi}{\partial p_2}\left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; 0, L^*_{22}\right) = Y^*_2$

(v) $\frac{\partial\pi}{\partial r_2}\left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; 0, L^*_{22}\right) = K^*_2$

(vi) $\frac{\partial\pi}{\partial w_{12}}\left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; 0, L^*_{22}\right) = L^*_{12}$

(vii) $\frac{\partial\pi}{\partial L^T_{11}}(p_1, r_1, w_{11}; L^{T*}_{11}, L^*_{21}) = w^T_{11}(*) + \frac{\partial w^T_{11}}{\partial L^T_{11}}(*)\, L^{T*}_{11} + C^1(L^{T*}_{11}) + \frac{1}{1+\rho}\frac{\partial w_{22}}{\partial L^T_{11}}(*)\, L^*_{22}$

(viii) $\frac{\partial\pi}{\partial L_{21}}(p_1, r_1, w_{11}; L^{T*}_{11}, L^*_{21}) = w_{21}(*) + \frac{\partial w_{21}}{\partial L_{21}}(*) L^*_{21} + \frac{\partial w^T_{11}}{\partial L_{21}}(*)\, L^{T*}_{11} + \frac{1}{1+\rho}\frac{\partial w_{22}}{\partial L_{21}}(*)\, L^*_{22}$

(ix) $\frac{\partial\pi}{\partial L_{22}}\left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; 0, L^*_{22}\right) = \frac{1}{1+\rho} w_{22}(*) + \frac{\partial w_{22}}{L_{22}}(*)\, L^*_{22}.$ [37/,38/]

The usual one period profit maximizing condition that marginal revenue equal marginal input costs for each factor are altered when intertemporal considerations enter. Thus, for example, training is supplied at a level such that the marginal 'revenue' of trainee labour in period 1 is equal to the discounted sum of the expected marginal input costs of trainee labour (equation (vii)).[39/,40/] Note, however, that the second period marginal input cost of (first period) trainee labour, $(\partial w_{22}(*)/\partial L^T_{11})L^*_{22}$, is negative.

We may rewrite equation (vii) as follows:

$$\frac{\partial\pi}{\partial L_{11}^{T}}(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^{*}) - \frac{1}{1+\rho}\frac{\partial w_{22}(*)}{\partial L_{11}^{T}} L_{22}^{*} = w_{11}^{T}(*) + \frac{\partial w_{11}^{T}(*)}{\partial L_{11}^{T}} L_{11}^{T*} + C^{1}(L_{11}^{T*}).$$

Thus, the marginal returns of "trainee" labour, where returns include both revenue from current production and (discounted) lower production costs in the future, are equal to the first period marginal input costs at the optimum activity levels. If the producer is a strong monopsonist and if there are low turnover rates in the labour market, he might expect the first period trainees to affect the future supply of "skilled" labour significantly, and the future return to training, in the form of $-\frac{1}{1+\rho}(\partial w_{22}/\partial L_{22}(*))\cdot L_{22}^{*}$, could be the major factor in his deliberations on how much training to supply. If trainees contribute nothing to current production, the partial derivative $\partial\pi(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^{*})/\partial L_{11}^{T}$ is equal to zero, and training will be provided only if the expected future returns warrant it. On the other hand, weak monopsony power and high turnover rates in the labour market might lead the producer to expect low future returns to training, in that $-\partial w_{22}/\partial L_{11}^{T}$ would assume small positive values, and the level of training provided would be determined to a large extent by the contribution made by trainees to current production. In the extreme case that $\partial w_{22}/\partial L_{11}^{T}=0$, there are no future returns at all, and the level of training will be determined by its first period costs and returns only.

Equation (viii) may be interpreted similarly. We emphasize at this point that in addition to the investment in training made by the producer, there may also be an investment involved in the hiring

of "skilled" workers in period 1. For we may rewrite equation (viii) as follows:

$$\frac{\partial \pi}{\partial L_{21}} (p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^{*}) - \frac{1}{1+\rho} \frac{\partial w_{22}}{\partial L_{21}} (*) L_{22}^{*} = w_{21}(*) + \frac{\partial w_{21}}{\partial L_{21}} (*) L_{21}^{*} + \frac{\partial w_{11}^{T}}{\partial L_{21}} (*) L_{11}^{T*}$$

Suppose that $\partial w_{22}/\partial L_{21} < 0$, i.e., that an increase in the first period "skilled" work force will increase the second period availability of "skilled" labour at a given wage rate. Then, once again possible decreased future production costs will influence the producer's decision as to the number of "skilled" workers to employ in the first period. He may hire more "skilled" workers in period 1 then is called for by first period profit maximization, in the expectation that they will lower his "skilled" labour costs in the future. Comments corresponding to those made above concerning the effects of the degree of monopsony power and labour market conditions may be made here as well.

If future "skilled" labour costs are expected to increase as a result of the use of more "skilled" workers in period 1, i.e., if $\partial w_{22}/\partial L_{21} > 0$, added future costs rather than returns are involved in the use of more "skilled" workers in period 1, and so the producer will probably try to substitute other factors of production for "skilled" labour.

The equations in II.11 may be transformed so as to correspond more closely to the form of the equilibrium conditions derived by Becker [2] to which we referred in the introduction. For this purpose, let us consider first the following equation, where w_{22}^{0} is a given positive constant and where the prices p_1, r_1, w_{11} are

given and suppressed in the notation:

II.12 $$w_{22}\ (L_{22}, L_{11}^{T}, L_{21}) = w_{22}^{o}.$$

This equation defines L_{22} implicitly as a function of L_{11}^{T}, L_{21} and w_{22}^{o}, the value of the function being precisely the amount of "skilled" labour that would be offered in period 2, given the wage rate w_{22}^{o} and the first period activities L_{11}^{T} and L_{21}. Implicit differentiation of II.12 first with respect to L_{11}^{T} and then with respect to L_{21}, shows

II.13 $$\frac{\partial L_{22}}{\partial L_{11}^{T}} = -\frac{\partial w_{22}}{\partial L_{11}^{T}} \Big/ \frac{\partial w_{22}}{\partial L_{22}}, \quad \frac{\partial L_{22}}{\partial L_{21}} = -\frac{\partial w_{22}}{\partial L_{21}} \Big/ \frac{\partial w_{22}}{\partial L_{22}}.$$

We make the following definitions:

II.14 $$\theta\ (w_{22}^{o}, L_{11}^{T}, L_{21}) \equiv \frac{\partial L_{22}}{\partial L_{11}^{T}} \quad ; \quad \phi\ (w_{22}^{o}, L_{11}^{T}, L_{21}) \equiv \frac{\partial L_{22}}{\partial L_{21}}.$$

Thus θ is the expected rate of change in the second period supply of "skilled" labour, at a given second period wage w_{22}^{o}, with respect to a small change in the first period supply of training. Similarly, ϕ is the expected rate of change of L_{22} with respect to a small change in the first period level of "skilled" labour employed.

If θ and ϕ are constant with respect to L_{11}^{T} and L_{21}, then integration of equations II.14 yields:

II.15 $$L_{22} = \theta\ (w_{22}^{o})\ L_{11}^{T} + \phi\ (w_{22}^{o})\ L_{21}.$$

Comparison with equation II.5 that θ and ϕ may be interpreted, after a suitable normalization of the units of measurement for "trainee" and "skilled" labour, as the retention rates of "trainee" and "skilled" labour respectively, at the second period wage rate w_{22}^{o}.[41/] In the more general case, the connection between θ, ϕ and the expected

retention rates of each type of labour is more complicated.[42/]

Now, let us replace equations (vii) and (viii) of II.11 with equations (vii)' and (viii)' obtained as follows: (vii)' is obtained by multiplying equation (ix) by $\theta\ (w_{22}(*), L_{11}^{T*}, L_{21}^{*})$ and adding the result to (vii); (viii)' is obtained by multiplying equation (ix) by $\phi\ (w_{22}(*), L_{11}^{T*}, L_{21}^{*})$ and adding the result to (viii). The results are:

$$\text{(vii)}'\quad \frac{\partial \pi}{\partial L_{11}^{T}} (p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^{*}) + \theta \cdot \frac{\partial \pi}{\partial L_{22}} \left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; o, L_{22}^{*}\right)$$

$$= [w_{11}^{T}(*) + \frac{\partial w_{11}^{T}(*)}{\partial L_{11}^{T}} L_{11}^{T*} + C^{1}(L_{11}^{T*})] + \theta \cdot \frac{w_{22}(*)}{1+\rho}$$

$$\text{(viii)}'\quad \frac{\partial \pi}{\partial L_{21}} (p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^{*}) + \phi \cdot \frac{\partial \pi}{\partial L_{22}} \left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; o, L_{22}^{*}\right)$$

$$= [w_{21}(*) + \frac{\partial w_{21}(*)}{\partial L_{21}} L_{21}^{*} + \frac{\partial w_{11}^{T}(*)}{\partial L_{21}} L_{11}^{T*}] + \phi \cdot \frac{w_{22}(*)}{1+\rho} .$$

Keeping in mind the above interpretations of θ and ϕ we see that $\theta \cdot \frac{\partial}{\partial L_{22}} \pi \left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; o, L_{22}^{*}\right)$ and $\phi \cdot \frac{\partial}{\partial L_{22}} \pi \left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; o, L_{22}^{*}\right)$ are the rates of change of the expected second period "revenue" with respect to changes in the first period level of training and employment of "skilled" labour respectively. On the other hand, the expressions in square brackets on the right hand side of each equation represent the marginal first period input costs for each type of labour. Second period anticipated discounted marginal input costs are $\theta \cdot w_{22}(*)/(1+\rho)$ and $\phi \cdot w_{22}(*)/(1+\rho)$. Thus equations (vii)' and (viii)' merely state the

equality between the expected discounted sum of marginal "revenues" and marginal input costs, for each factor in turn, which is the familiar equilibrium condition for the maximization of a discounted sum of profits.

A measure of the anticipated future return to the producer from training is given by the expression
$\theta \cdot \partial/\partial L_{22}\ \pi\ (p_2/1+\rho,\ r_2/1+\rho,\ w_{12}/1+\rho;\ o,\ L_{22}^{*}) - \theta \cdot w_{22}(*)/1+\rho$, which, after using equation (ix) of II.11, may be rewritten as the product
$\theta \cdot [\frac{1}{1+\rho}\ \frac{\partial w_{22}(*)}{\partial L_{22}}\ L_{22}^{*}]$. Hence the future return depends on the degree of specificity of the training, as reflected by the anticipated extent of the monopsony power over "skilled" labour, i.e. on $(\partial w_{22}(*)/\partial L_{22})L_{22}^{*}$, and on external market conditions as they affect the anticipated future actions of trainees, i.e., on θ. If future returns are non-zero, the producer will deviate from first period profit maximization, and be willing to incur the first period costs reflected by the excess of first period marginal input costs over first period marginal revenue, i.e., by the fact that $\frac{\partial}{\partial L_{11}^{T}}\pi\ (p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^{*}) - [w_{11}^{T}(\ ) + \frac{\partial w_{11}^{T}}{\partial L_{11}^{T}}(*)L_{11}^{T*} + C^{1}(L_{11}^{T*})] < o.$/[43]
Similar statements can be made concerning equation (viii)' and the "investment" in "skilled" labour.

We note, however, that though the form of equations (vii)' and (viii)' is appealing in that it conforms closely to the standard formulation of the equilibrium conditions and thus permits simple interpretations, it may be inferior for empirical work. This we shall comment on in section III. [illegible] of our

Before doing so, we remark that our model may be readily altered so as to model the supply of general training. For suppose that "skilled" workers have the option of working in several other firms and industries where their productivity is the same as in the firm providing the training. In a perfectly competitive market, the producer will then be a price taker with respect to "trainee" and "skilled" labour, and he will not expect his future supply of "skilled" labour to be influenced by his current activities. Assuming a long run equilibrium condition of zero profits and using a result proved by Diewert ([14 ; Lemma 4.15]), we derive that the optimum values of L_{11}^T, L_{21} and L_{22} are defined by the following equations:[44/]

$$\text{(vii)"} \quad \frac{\partial\pi}{\partial L_{11}^T}(p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^*) = w_{11}^{To} + C'(L_{11}^{T*})$$

$$\text{(viii)"} \quad \frac{\partial\pi}{\partial L_{21}}(p_1, r_1, w_{11}; L_{11}^{T*}, L_{21}^*) = w_{21}^o$$

$$\text{(ix)"} \quad \frac{\partial\pi}{\partial L_{22}}\left(\frac{p_2}{1+\rho}, \frac{r_2}{1+\rho}, \frac{w_{12}}{1+\rho}; o, L_{22}^*\right) = \frac{1}{1+\rho} w_{22}^o ,$$

where w_{11}^{To} and w_{21}^o are the market determined wage rates for "trainee" and "skilled" labour respectively, w_{22}^o is the "skilled" wage rate expected to prevail in the second period, and where $p_1, r_1, w_{11}, p_2, r_2, w_{12}$ and ρ are as before. These equations, combined with equations (i) to (vi) of II.11 completely describe the producer's behaviour. We see that intertemporal considerations have disappeared and that producers will supply training only if first period costs and returns warrant it. If wage rates for trainees and direct training costs are high, or if the technology is such that the productivity of trainees is low, there may not exist a solution to the above three equations.

In that case, no general training at all is provided, and equations (viii)" and (ix)" with L_{11}^{T*} set equal to zero, define the optimum levels of "skilled" labour in each period.

Finally, before proceeding to discuss the applicability of our model for empirical work, we indicate how the model may be extended to a framework of several time periods. Let us suppose then that the producer faces a time horizon of τ periods. He faces first period supply curves for "skilled" labour and "trainee" labour described by II.1 and II.2 respectively. As in II.3, the producer forms expectations of the prices of capital, "unskilled" labour and the output good which will prevail in each of the future periods, and of the discount rates which will prevail between each of the periods.

Following the reasoning preceeding II.4, we could describe the supply curve of "skilled" labour facing the firm in period t by the expected (inverted) supply function:

I.16 $$w_{2t}=w_{2t}\ (L_{2t},L_{11}^{T},\ldots,L_{1,t-1}^{T},L_{21},\ldots,L_{2,t-1};\ p_1,r_1,w_{11})\quad,\quad t=2,3,\ldots,\tau;$$

The supply curve in period t would then be affected by the training supplied, and by the number of "skilled" workers employed, in all of the preceeding periods. However, in order to keep the model sufficiently simple so that it may be applied empirically, we make the approximation that the supply curve in any period is affected only by the amounts of "trainee" and "skilled" labour used in the immediately preceeding period; that is, we assume expected (inverted) supply functions of "skilled" labour of the form:

II.17 $$w_{2t}=w_{2t}\ (L_{2t},L^{T}_{1,t-1},L_{2,t-1};\ p_1,r_1,w_{11}) \quad , \quad t=2,3,\ldots,\tau.$$

Trainees from period (t-2) say, still affect the t^{-th} period supply of "skilled" labour, only indirectly, by increasing the availability, and hence probably the employment of, "skilled" labour in period (t-1). Their direct effect on the t^{-th} period supply to a firm, say in the form of graduates that leave the firm at the end of period (t-2), and return as "skilled" workers only at the start of period t, is probably of little significance in comparison both with the effects of $L^{T}_{1,t-1}$, and $L_{2,t-1}$ on the t^{-th} period "skilled" supply curve, and with the effect which training in period (t-2) has on the supply curve of "skilled" labour in the succeeding period, (t-1). Thus, II.17 is a reasonable approximation to II.16, and the results below are valid.

Lastly, we may consider the anticipated direct costs of training and the anticipated supply of "trainee" labour in each of the future periods. We assume that neither is affected by the prior activities of the firm.

Proceeding precisely as above, we assume that the producer strives to maximize the expected discounted sum of profits over the τ periods, and derive a set of equilibrium equations. It is easy to see that equations (i) through (viii) from II.11 are contained in this new set of equations. Equation (ix) is altered, since training is now conceivably carried on in period 2 and since the number of "skilled" workers employed in period 2, will affect the third period supply curve of "skilled" labour. In addition, there will be many more equations defining other optimal planned activities such as L^{T*}_{1t} and L^{*}_{2t} , $t=3,4,\ldots,\tau$. However, in an empirical analysis, equations (i) through (viii) are the most important. This we show in the following section.

III. ECONOMETRICS

The set of equations II.11 are, theoretically at least, well suited to econometric analysis. Substitution of the appropriate version of Diewert's functional form for the variable profit function yields a system of equations linear in the unknown parameters, and so linear regression techniques may be applied to estimate them. (See I.5 and the discussion following it.) This assumes, of course, knowledge, of the price expectation functions p_2, r_2, w_{12} and ρ (II.3), the supply functions w_{21} (II.1) and w_{11}^T (II.2), the expected supply function w_{22} (II.4), and the training direct cost function $C(L_{11}^T)$. In addition, we must have data on first period prices and activities and on second period plans, and then the nine equations in II.11 will enable us to estimate the variable profit function and hence the underlying technology.[45/]

The above information is rarely known completely however, and alternate procedures may have to be followed. For example, if one is not certain of price expectations, or of the second period plans of the producer, with the exception of L_{22}^*, equations (i), (ii), (iii), (vii) and (viii) alone may be used to estimate the profit function. In this case, the costs of incomplete information are a loss of four degrees of freedom for each observation. (We note that the use of equations (vii)' and (viii)' requires knowledge of price expectations.)

If the supply functions w_{11}^T, w_{21} and w_{22} are not known, we may postulate functional forms for them which are linear in the unknown parameters. Substitution of these forms and their partial

derivatives into our equations again yields a system of equations which is amenable to the techniques of linear regression analysis. We note that this is not the case if equations (vii) and (viii) are replaced by (vii)' and (viii)', which is the primary reason that we claimed that the latter two are inferior for empirical work. This linearity in equation II.11 which allows us to estimate the supply functions is also the principal reason for our having chosen the variable profit function, rather than the profit function, to represent the technology. (See footnote 30). Note however, that we must still have an estimate of L_{22}^{*}, the planned future use of "skilled" labour. If economic conditions are fairly stable, or if we can otherwise attribute good foresight to the producer, it may not be a poor approximation to substitute for the planned future employment of "skilled" labour, the actual future employment, for which presumably data are available.

Following the form of the expected supply function w_{22} defined in II.5, an example of a functional form we might wish to postulate for w_{22} is the one defined implicitly by

III.1 $$L_{22} = w_{22}^{\frac{1}{2}}[aL_{11}^{T} + b\,L_{21}],$$

where a and "b" are non-negative parameters to be estimated, with b/a representing the ratio of the retention rate of "skilled" workers to the retention rate of trainees. Unfortunately, when we solve for w_{22} above, the parameters do not appear linearly and so their estimation is made difficult.

An alternative is a function of the form

III.2 $$w_{22} = L_{22}^2 \left[\frac{a}{L_{11}^T} + \frac{b}{L_{11}^{T\frac{1}{2}} L_{21}^{\frac{1}{2}}} + \frac{c}{L_{21}}\right],$$

where a, b, and c are nonnegative parameters to be estimated. They do appear linearly and so linear regression techniques may be applied. Certainly, more general forms may be found. However, we will not proceed further into postulating functional forms for w_{22}, and for w_{11}^T and w_{21}, but leave it to the reader who wishes to perform some empirical work with the model to do so.

The discussion to this point has been concerned with an empirical analysis in a two period world. Since real firms do not operate in a two period framework, most empirical work will require the application of the extended model described at the end of the last section. Suppose, for example, that we have time series data on the prices and quantities of the factors employed and the goods produced by a firm. We take τ, the planning horizon, to be longer than the sample period. Suppose that there existed reasonably constant external economic conditions over the sample period, so that it is reasonable to assume that the producer faced the same (inverted) supply curves for "skilled" and trainee labour throughout the sample period, that his expected supply function II.17 was also constant, and that the "skilled" labour planned in period t for use in period t+1, was actually employed in period t+1. Lastly, suppose that we have estimates of the expected discount rate between each of the periods in the sample period, and that we know the direct cost of training function which prevailed in each period. Then, equations (i), (ii), (iii), (vii) and (viii) from II.11, with each year, say, of the sample period, taken in turn to be the first period of the planning horizon,

may be used to estimate the underlying technology, (also assumed to have remained constant over the sample period), as well as the supply curves for "skilled" and "trainee" labour, and the producer's expectation function II.17.

Footnotes

1. I am indebted to Erwin Diewert of the University of British Columbia and to Ken Scott of the Research Branch, Department of Manpower and Immigration, for suggesting the topic to me and for some helpful comments.

2. Good discussions of functional form specification may be found in [7] and [11]. Some of the new functional forms which have been developed are: the Generalized Leontief Function by Diewert [8], the Cresh Production Function by Hanoch [17], the translog function by Christensen, Jorgensen and Lau [6], and an as yet unnamed function by Denny [7]. In [11], Diewert has specified some forms for the cases of multiple outputs and some fixed inputs. It is left to the reader to investigate the advantages specific to each of these forms. We mention only the two most important properties common to most of them: firstly, they are linear in the unknown parameters so that linear regression techniques may readily be applied to the task of estimating the unknown parameters; and secondly, they contain precisely the number of parameters needed in order to provide a second order approximation to an arbitrary production function satisfying the appropriate regularity conditions. Since statements on first and second order derivatives correspond to statements in elasticities of substitution, the parameters may be chosen so that the elasticities of substitution assume any predetermined set of values.

3. The reader is asked to refer to the references mentioned in the preceding footnote for a fuller discussion of some of the statements made here. In the next section of the paper we shall dwell a little more on those results from the theory of duality which we shall require in our model.

4. Thus, for example, in [4], the authors, by testing hypotheses about elasticities of substitution, are able to test for the existence of a consistent aggregate index of labour inputs; while in [5], similar hypothesis testing is used in seeking an economic interpretation for the concept of the energy real value added ratio.

5. In spite of the apparent intertemporal framework in the model in [13], producers are maximizing short run profits only.

6. In Canada, the Department of Manpower and Immigration, through the Canada Manpower Industrial Training Program, is engaged in the subsidization of on-the-job training. An accurate evaluation of the program calls for an understanding of producers' behaviour in the absence of subsidization, and of the effects of varying levels of subsidization on their behaviour. It is hoped that the model developed in this paper may contribute to such an understanding.

7. See [23] for example. The same technique has also been applied in models of the demand for training where it is assumed that individuals wish to maximize the discounted integral of future earnings ([3]), or the discounted integral of a utility function depending on earnings and other variables ([27]).

8. The study by Millar [23] is such an analysis.

9. By monopsony power over a factor we shall mean that the equilibrium wage for the factor is a decision variable for the firm, at least over some interval, and that the supply of labour to the firm is not a totally elastic function of the equilibrium wage in that interval. That our assumption is not a restrictive one in the context of specific training will be argued below.

10. Some of the weaknesses of the model that are not so easily removed include: (a) the assumption of certain expectations, i.e., we are neglecting the phenomena of risk and uncertainty; (b) we neglect the problem of the optimum distribution of total training times among employees, i.e., is full-time training of fewer workers preferred by the producer to part-time training of more workers, and so on; (c) the training technology - the output of "skilled" workers relative to the resources devoted to their training - does not receive much attention.

11. An excellent survey of duality and its uses in economics may be found in [14].

12. We choose the variable profit function, rather than the profit function, to represent technology, the reasons for which will become clear below.

13. T is a convex set if $z' \in T$, $z'' \in T$, $o \leq \lambda \leq 1$, implies $\lambda z'+(1-\lambda)\ z'' \in T$.

14. T is a cone if $z \in T, \lambda \geq 0$ implies that $\lambda z \in T$.

15. t is a continuous from above function if $t(z_n) \geq t(z_o)$, n=1,2,---, and $\lim_{n\to\infty} z_n = z_o$, implies that $\lim_{n\to\infty} t(z_n) = t(z_o)$.

16. The function t is a proper concave function over a convex set S if (a) for every $z' \varepsilon S$, $z' \varepsilon S$ and $0 \leq \lambda \leq 1$, we have $t(\lambda z' + (1-\lambda) z'') \geq \lambda t(z') + (1-\lambda)\, t(z'')$; (b) $t(z) < \infty$ for every $z \varepsilon S$ and (c) $t(z) > -\infty$ for at least one $z \varepsilon S$.

17. The function π is a convex function in p, for all p in a convex set S if $-\pi$ satisfies the appropriate version of condition (a) in footnote 16.

18. The corresponding lemma for profit functions is due to Hotelling [18]. Both versions are proved by Diewert in [11].

19. To be more pricise, Diewert shows that if the parameters are such that π satisfies the regularity conditions over some non empty set of vectors $(p;v)$, then there exists a variable profit function π', and a closed convex set S, such that $\pi(p;v) = \pi'(p;v)$ for all $(p;v)$ in S.

20. For example, the Allen [1] Uzawa [28] Partial Elasticity of Transformation between goods i and h, at a given price, variable input and output vector (p;v), may be defined as $\sigma_{ih} = -\pi_{ih}\,\pi \big/ \pi_i\,\pi_h$, where $\pi_{ih} \equiv \partial^2\pi(p;v) \big/ \partial p_i\,\partial p_h$, and π_i and π_h are the appropriate first order partial derivatives. In view of the Modified Hotelling Lemma, we see that σ_{ih} is a normalization of the response of the variable profit maximizing output (or input) i to a change in the price of output (or input) j.

21. Differentiation of I.5 yields:

$$\pi_{ih} = -\tfrac{1}{2}\, a_{ih}\, (\tfrac{1}{2}\, p_i^2 + \tfrac{1}{2}\, p_h^2)^{-3/2}\, p_i p_h (\sum_{j=1}^{J} xj).$$

Using the definition of the Allen -Uzawa Partial Elasticity of Transformation between goods i and h,σ_{ih}, we see that σ_{ih}=0 if and only if α_{ih}=0.

22. From I.5 we may derive easily that

$$\pi^{jk} \equiv \frac{\partial^2 \pi\ (p;v)}{\partial v_j\ \partial v_k} = \tfrac{1}{2}\, b_{jk}\, (x_j x_k)^{-\frac{1}{2}}\, (\sum_{i=1}^{L} p_i).$$

Of course, π^{jk} is the rate of change in the marginal variable profit of of the j^{th} fixed input (or output) with respect to a change in the level of the k^{th} fixed input (or output), and π^{jk}=o if and only if b_{jk}=0.

23. In [15] Fuss has used duality theory to model long run optimization with respect to fixed inputs, which allows for flexibility in the choice of short run inputs and outputs. Long run optimization may take into account the various types of training and hence different training costs involved in the job definitions corresponding to the various feasible technologies. This is discussed by Scoville in [25].

24. Since in our model the roles of inputs and outputs do not change, we index both inputs and outputs positively. The results in section I. may be readily translated in terms of the new sign convention.

25. The assumption of constant returns to scale may be dropped. All of the results proved in I. have analogues in the case of decreasing returns to scale in production, and so the model may be extended to handle that type of technology. The extension to increasing returns to scale is a little more difficult. Nevertheless if there are increasing returns to scale in all four factors, but decreasing returns to scale with respect to capital and "unskilled" labour when "trainee" and "skilled" labour inputs are fixed, and if the producer faces sufficiently rising expected input costs so that the maximum profit he can hope to earn over the two periods is finite, (see the discussion below following II.8), then the producer's optimal activities may be derived as below.

26. We have argued that the first period supply of "skilled" labour to the firm, L_{21}, is a monotonic increasing function of the wage rate offered w_{21}. The fünction may thus be inverted to express w_{21} as a function of L_{21}.

27. Not all of these conditions are essential. However, we do not concern ourselves with specifying a minimal set of regularity conditions on w_{21} and the other functions defined below.

28. This function may be defined only for a subset of the possible values of its arguments, i.e., for those values consistent with an equilibrium situation.

29. Assuming that workers entertain expectations of future prices which are functions of current prices, possibly distinct from the producer's expectations, (see II.3 below), the functional forms in II.1 and II.2 are sufficiently general to reflect intertemporal utility maximization on their part.

30. Our model may be readily extended to cover the case of monopoly power in production. If, in addition, there is monopsony power in all factor inputs, then the technique mentioned by Diewert ([12], footnote 18) may be adopted. (The analysis of the perfectly competitive case is duplicated, except that the price of each factor is replaced in the demand equations by its marginal input costs.) The difficulty is, however, that all the supply functions (of inputs) and demand functions (for outputs) faced by the producer, must be known in advance, at least around the optimum points. Our technique, which may be used in the event of price taking behaviour with respect to at least one commodity, allows one to apply linear regression techniques to estimate these functions. This will be commented on further below.

31. Price expectations, as well as the supply functions above, depend also upon prices and activities in the rest of the economy. Ours is thus a partial analysis in that it assumes a given set of "external" economic conditions. Note also that the wage rates w_{11}^T and w_{21} are decision variables for the producer and so need not enter directly into his expectation functions.

32. This expectation may derive in part from contractual arrangements between the producer and trainees.

33. If α and β are any non-negative bounded functions, an appropriate choice of units of measurement for "trainee" and "skilled" labour will allow us to assume that α and β assume values in the unit interval.

34. Implicit differentiation of II.5 yields the following:

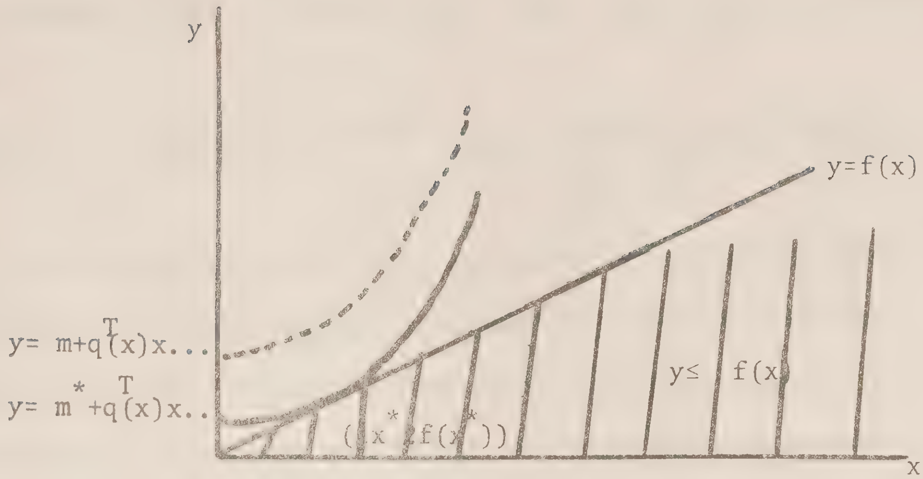
$$\frac{\partial w_{22}}{\partial L_{11}^T} = -\left(\alpha + \frac{\partial \alpha}{\partial L_{11}^T} L_{11}^T\right) \Big/ \left(\frac{\partial \alpha}{\partial w_{22}} L_{11}^T + \frac{\partial \beta}{\partial w_{22}} L_{21}\right),$$

$$\frac{\partial w_{22}}{\partial L_{21}} = -\left(\beta + \frac{\partial \beta}{\partial L_{21}} L_{21}\right) \Big/ \left(\frac{\partial \alpha}{\partial w_{22}} L_{11}^T + \frac{\partial \beta}{\partial w_{22}} L_{21}\right),$$

$$\frac{\partial w_{22}}{\partial L_{22}} = 1 \Big/ \left(\frac{\partial \alpha}{\partial w_{22}} L_{11}^T + \frac{\partial \beta}{\partial w_{22}} L_{21}\right).$$

35. We note that since the producer's expectations for period 2 may not be fulfilled, he may find later that his chosen first period activities were not optimal and that he did not succeed in maximizing the actual discounted sum of profits.

36. Maximization problem II.8 is of the form $\max_{x \geq 0} f(x) - q(x)^T x$, where is a linear homogeneous concave production function, $x = (x_1, x_2, x_3)$ is a vector of inputs, and where $q(x) = (q_1(x), q_2(x))$ is the vector of (inverted) supply functions facing the producer. This problem in turn may be rewritten as $\max_{\substack{y \leq f(x) \\ x \geq 0}} y - q(x)^T x$, the solution of which, assuming that input costs are sufficiently increasing, be represented diagramatically as follows:



The optimal point $(x^*, f(x^*))$ is the point of tangency between the production line $y = f(x)$, and the convex isσ profit curves $y = m + q(x)^T x$. Strict convexity of $q(x)^T x$ assures that this point of tangency is unique. The maximum profit attained is m^*.

37. We have made the obvious abbreviation in notation; e.g., $w_{21}(*)$ stands for $w_{21}(L_{21}^*; p_1, r_1, w_{11})$, and so on.

38. If we wish to determine the effects of government subsidization of training which takes the form, for example, of sharing the direct training costs incurred by the producer, we replace the direct cost function in (vii) by $(1-\lambda)\ C(L_{11}^T)$, where λ is the government's share of the costs. The new set of equations will define implicitly the new optimum activities, including the new optimum level of training supplied. Similarly, if subsidization takes the form of wage sharing, the supply function $w_{11}{}^T$ may be replaced by $(1-\mu)\ w_{11}{}^T$, where μ is the share paid by the government. The effects of government subsidization on the supply of general training may be determined similarly from some equations presented below. (See equations (vii)", (viii)" and (ix)").

39. We are assuming, of course, that a positive amount of training is provided at the optimum point. A necessary and sufficient condition for this to be true is that

$$\frac{\partial\pi}{\partial L_{11}^T}(p_1,r_1,w_{11};0,L_{21}^o) > w_{11}^T(0,L_{21}^o) + C'(0) + \frac{1}{1+\rho}\frac{\partial w_{22}}{\partial L_{11}^T}(L_{22}^o,o,L_{21}^o,)\ L_{21}^o,$$

where L_{21}^o and L_{22}^o are the profit maximizing employment of "skilled" labour in the first and second periods respectively, subject to the constraint that no training is provided in period 1 (and hence in either period); i.e., they are the solutions to equations (viii) and (ix) when L_{11}^{T*} is set equal to zero.

40. The variable profit function gives the value added by "trainee" and "skilled" labour, to which we refer as "revenue".

41. To avoid the ambiguity caused by negative retention rates, we assume here and in the discussion to follow that $\partial w_{22}/\partial L_{21} \geq o$, and hence that $\phi \geq o$.

42. Differentiation of II.5 shows that

$$\frac{\partial L_2}{\partial L_{11}^T} = \alpha + \frac{\partial\alpha}{\partial L_{11}^T}L_{11}^T \quad \text{and} \quad \frac{\partial L_{22}}{\partial L_{21}} = \beta + \frac{\partial\beta}{\partial L_{21}}L_{21},$$

where α and β are the retention rates. Hence we see that for the expected supply function defined by II.5, we have

$$\Theta = \alpha + \frac{\partial\alpha}{\partial L_{11}^T}L_{11}^T \quad \text{and} \quad \phi = \beta + \frac{\partial\beta}{\partial L_{21}}L_{21}.$$

43. In the literature reference is often made to the costs of foregone production involved in the supply of training. As we see here, these costs are more accurately labelled costs of foregone profit.

The costs of supervision referred to in the introduction are reflected in our model by the fact that the productivity of "skilled" workers may be adversely affected by the presence of trainees, so that the optimum number of "skilled" workers L_{21}^*, is not optimum in the absence of training; i.e., we probably have that

$$\frac{\partial}{\partial L_{21}} \pi\,(p_1,r_1,w_{11};0,L_{21}^*) + \phi(w_{22}(*),0,L_{21}^*)\,\frac{\partial}{\partial L_{21}}\,\pi\left(\frac{p_2}{1+\rho},\frac{r_2}{1+\rho},\frac{w_{12}}{1+\rho};0,L_{22}^*\right)$$

$$\neq\ [w_{21}(*) + \frac{\partial w_{21}(*)}{\partial L_{21}}\,L_{21}^*] + \phi(w_{22}(*),0,L_{21}^*)\,\frac{w_{22}(*)}{1+\rho}\ .$$

The costs of training resulting from increased wear and tear on machines could also be included in our model if we dealt with the formation of capital as did Diewert in [13], and allowed for the effects of the number of trainee workers on the depreciation rates of capital.

44. In the case of decreasing returns to scale in production, the profit which the producer expects to earn is bounded, and the equations below follow from profit maximization. With constant returns to scale in production and the assumption zero profits, Diewert shows that if L_{11}^{T*}, L_{21}^* and L_{22}^* are chosen so as to minimize the cost of producing the variable profits sum $\pi(p_1,r_1,w_{11};L_{11}^{T*},L_{21}^*) + \pi\left(\frac{p_2}{1+\rho},\frac{r_2}{1+\rho},\frac{w_{12}}{1+\rho};0,L_{22}^*\right)$,

then they satisfy the equations given.

45. Since ours is a partial analysis, allowance will have to be made in an empirical analysis for changing external economic conditions and their effects on, for example, the producer's expectation functions. These effects would probably be significant if we are using time series data covering a period in which general economic conditions changed significantly.

References

1. Allen, R.G.D., Mathematical Analysis for Economists, Macmillan, 1938.

2. Becker, G., Human Capital, New York: N.B.E.R., 1964.

3. Ben-Porath, Y., "The Production of Human Capital and the Life of Earnings", Journal of Political Economy, 75, August 1967, pp. 352-65.

4. Berndt, E.R. and Christensen, L.R., "Testing for the Existence of a Consistent Aggregate Index of Labour Inputs", Paper No.79, Office of Emergency Preparedness, Washington, D.C., August 1972.

5. Berndt, E.R. and Wood, D.O., "The Derived Demand for Energy Inputs", Paper Presented at the Meetings of the Econometric Society, Toronto, Canada, December 1972.

6. Christensen, L.R., Jorgensen, D.W. and Lau, L.J., "Conjugate Duality and the Transcendental Logarithmic Production Function", Econometrica, Vol.39, No.4, July 1971, pp. 255-56.

7. Denny, M., "The Relationship Between Functional Forms for the Production System", Working Paper No. 7007, Institute for Quantitative Analysis, University of Toronto, 1970.

8. Diewert, W.E., "An Application of the Shepherd Duality Theorem: A Generalized Leontief Production Function," Journal of Political Economy, Vol.79, May/June 1971, pp. 481-507.

9. __________, "Canadian Labour Markets; A Neoclassical Econometric Approach", Research Branch, Department of Manpower and Immigration, Ottawa, 1971.

10. __________, "Choice on Labour Markets and the Theory of the Allocation of Time", Research Branch, Department of Manpower and Immigration, Ottawa, 1971.

11. __________, "Functional Forms for Profit and Transformation Functions", University of British Columbia and Research Branch, Department of Manpower and Immigration, Ottawa, 1972.(To appear also in J. ECO. THEORY).

12. __________, "Functional Forms for Revenue and Factor Requirements Functions", University of British Columbia and Research Branch, Department of Manpower and Immigration, Ottawa, 1972. (To appear also in INT.ECO.REVIEW.)

13. __________, "Walras Theory of Capital Formation and the Existence of a Temporary Equilibrium", University of British Columbia and Research Branch, Department of Manpower and Immigration, Ottawa, 1972.

14. __________, "Applications of Duality Theory", Discussion Paper No. 89, University of British Columbia. Presented at the Meetings of the Econometric Society, Toronto, Canada, December 1972.

15. Fuss, M., "The Structure of Technology Over Time: A Model for Testing the 'Putty-Clay' Hypothesis", Discussion Paper No. 141, Harvard Institute of Economic Research, Cambridge, Massachusetts, April 1972.

16. Gorman, W.M., "Measuring the Quantities of Fixed Factors", in Value, Capital and Growth: Papers in Honour of Sir John Hicks, J.N. Wolfe (ed.), Aldine Publishing Co., 1968, pp. 141-172

17. Hanoch, G., "Generation of New Production Functions Through Duality," Discussion Paper No. 118, Harvard Institute of Economic Research, Cambridge, Massachusetts, April 1970.

18. Hotelling, H., "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions", Journal of Political Economics, Vol.40, 1932, pp. 557-616.

19. Johnson, T., "Returns from Investment in Human Capital", American Economic Review, September 1970, pp. 546-560.

20. Jorgensen, D.W. and Lau, L.J., "The Duality of Technology and Economic Behaviour in Production Theory", Discussion Paper No. 252, Harvard Institute of Economic Research, Cambridge, Massachusetts, August 1972.

21. McFadden, D., "Cost, Revenue and Profit Functions: A Cursory Review", Working Paper No. 86, IBER, University of California at Berkley, March 1966.

22. __________, "Cost, Revenue and Profit Functions", in An Econometric Approach to Production Theory, edited by D. McFadden, forthcoming, (preliminary draft, April 8, 1970).

23. Millar, J.R., " Theory of On-The-Job Training", Working Paper 71-19, School of Business, University of Toronto, June 1971.

24. Mincer, J., "On-The-Job Training: Costs, Returns and Implications", Journal of Political Economy, October 1962, pp. 50-79.

25. Scoville, J.G., "A Theory of Jobs and Training", Discussion Paper No. 68, Harvard Institute of Economic Research, Cambridge, Massachusetts, March 1969.

26. Shephard, R.W., Cost and Production Functions, Princeton University Press, 1953.

27. Stafford, F.P. and Stephan, P.E., "Labour, Leisure, and Training Over the Life Cycle", Paper Presented at the Meetings of the Econometric Society, Toronto, Canada, December 1972.

28. Uzawa, H., "Production Functions with Constant Elasticities of Substitution", Review of Economic Studies, Vol. 29, 1962, pp. 291-99.